# DAFTAR ISI

HALAMAN SAMPUL i

HALAMAN JUDUL ii

HALAMAN PERSETUJUAN DOSEN PEMBIMBING iii

HALAMAN PERNYATAAN DOSEN PENGUJI iv

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI vi

ABSTRAK viii

KATA PENGANTAR ix

DAFTAR ISI xi

DAFTAR TABEL xv

DAFTAR GAMBAR xviii

DAFTAR LAMPIRAN xx

BAB 1 PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Rumusan Masalah 3

1.3 Ruang Lingkup 4

1.4 Tujuan dan Manfaat 4

1.4.1 Tujuan 4

1.4.2 Manfaat 5

1.5 Metode Penelitian 5

1.5.1 Metode Pengumpulan Data 5

1.5.2 Metode Perancangan Sistem 6

1.5.3 Metode Evaluasi 7

1.6 Sistematika Penulisan 7

BAB 2 TINJAUAN REFERENSI 9

2.1 *Artificial Intelligence* 9

2.2 *Natural Language Processing* (NLP) 9

2.3 *Machine Learning* 11

2.3.1 Pengertian *Machine Learning* 11

2.3.2 Algoritma *Machine Learning* 12

2.3.2.1 *Support Vector Machine* 12

2.3.2.2 Naïve Bayes 14

2.3.2.3 *Logistic Regression* 15

2.3.2.4 *Gradient Boosting* 17

2.3.2.5 *Linear Discriminant Analysis* 17

2.3.3 *Library Machine Learning* 17

2.3.3.1 Scikit-learn 17

2.3.3.2 NumPy 18

2.3.3.3 Pandas 18

2.3.3.4 Imbalanced-Learn 18

2.3.3.5 NLTK (*Natural Language Toolkit*) 18

2.4 *Deep Learning* 19

2.4.1 Pengertian *Deep Learning* 19

2.4.2 Arsitektur *Deep Learning* 20

2.4.2.1 MLP (*Multilayer Perceptron*) 20

2.4.2.2 CNN (*Convolutional Neural Network*) 21

2.4.2.3 LSTM (*Long Short-Term Memory*) 22

2.4.2.4 GRU (*Gated Recurrent Unit*) 23

2.4.3 *Library Deep Learning* 24

2.4.3.1 Keras 24

2.4.3.2 Theano 24

2.5 *Resampling* 24

2.6 Python 25

2.7 Eclipse IDE 26

2.8 XAMPP 26

2.9 Apache HTTP Server 26

2.10 *Web Service* 26

2.11 Facebook 29

2.12 my*Personality* 30

2.13 *Apply Magic Sauce* 31

2.14 *The Big Five Traits* 31

2.15 LIWC 33

2.16 SPLICE 35

2.17 Penelitian Terkait 36

2.17.1 *Personality and Patterns of Facebook Usage* 36

2.17.2 *Recognising Personality Traits Using Facebook Status Updates* 40

2.17.3 *Personality, Gender, and Age in the Language of Social Media:*

*The Open-Vocabulary Approach* 44

2.17.4 Sistem Prediksi Kepribadian “*The Big Five Traits*” Dari Data

Twitter 50

2.17.5 *Personality Prediction Based on Twitter Information in Bahasa* 52

2.17.6 *Personality Traits Recognition on Social Network - Facebook* 57

2.17.7 *Deep Learning-Based Document Modeling for Personality Detection*

*from Text* 60

BAB 3 METODOLOGI 63

3.1 Analisis Masalah 63

3.2 Usulan Pemecahan Masalah 64

3.3 Kerangka Berpikir 65

3.3.1 *Dataset Collecting* 65

3.3.2 *Data Preprocessing* 70

3.3.3 *Feature Selection* 75

3.3.4 *Model Learning Process* 89

3.3.5 *Prediction Model Testing* 92

3.3.6 *Final Result of Prediction Based on Big Five Model* 93

3.4 Rancangan Layar 93

3.4.1 Rancangan Layar *Homepage* 93

3.4.2 Rancangan Layar *Privacy Policy* 94

3.4.3 Rancangan Layar *Result* 95

BAB 4 HASIL PENELITIAN 97

4.1 *Testing Environment* 97

4.1.1 Spesifikasi Sistem 97

4.1.1.1 *Device* 97

4.1.1.2 *Server* 97

4.1.2 Perangkat Lunak 97

4.1.3 *Library* 97

4.2 Hasil 98

4.2.1 Distribusi Data 98

4.2.2 Faktor Skenario Percobaan 99

4.2.2.1 *Feature Selection* 99

4.2.2.2 *Resampling* 99

4.2.3 Implementasi *Machine Learning* 100

4.2.3.1 Skenario 100

4.2.3.2 Dataset my*Personality* 101

4.2.3.2.1 Skenario 1 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 102

4.2.3.2.2 Skenario 2 (Percobaan dengan menggunakan *Features*

*Selection*) 103

4.2.3.2.3 Skenario 3 (Percobaan dengan menggunakan

*Undersampling*) 105

4.2.3.2.4 Skenario 4 (Percobaan dengan menggunakan

*Oversampling*) 107

4.2.3.2.5 Skenario 5 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 109

4.2.3.2.6 Skenario 6 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 111

4.2.3.2.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset* my*Personality* 113

4.2.3.3 Dataset Manual Gathering 114

4.2.3.3.1 Skenario 7 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 114

4.2.3.3.2 Skenario 8 (Percobaan dengan menggunakan *Features*

*Selection*) 115

4.2.3.3.3 Skenario 9 (Percobaan dengan menggunakan

*Undersampling*) 117

4.2.3.3.4 Skenario 10 (Percobaan dengan menggunakan

*Oversampling*) 118

4.2.3.3.5 Skenario 11 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 119

4.2.3.3.6 Skenario 12 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 121

4.2.3.3.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset Manual Gathering* 122

4.2.3.4 Dataset Gabungan 123

4.2.3.4.1 Skenario 13 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 124

4.2.3.4.2 Skenario 14 (Percobaan dengan menggunakan *Features*

*Selection*) 125

4.2.3.4.3 Skenario 15 (Percobaan dengan menggunakan

*Undersampling*) 127

4.2.3.4.4 Skenario 16 (Percobaan dengan menggunakan

*Oversampling*) 128

4.2.3.4.5 Skenario 17 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 129

4.2.3.4.6 Skenario 18 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 131

4.2.3.4.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset* Gabungan 132

4.2.3.5 Kesimpulan Implementasi *Machine Learning* 133

4.2.4 Implementasi *Deep Learning* 135

4.2.4.1 Skenario 135

4.2.4.2 *Dataset* my*Personality* 137

4.2.4.2.1 Skenario 1 (Percobaan tanpa menggunakan

*Resampling*) 137

4.2.4.2.2 Skenario 2 (Percobaan dengan menggunakan

*Undersampling*) 138

4.2.4.2.3 Skenario 3 (Percobaan dengan menggunakan

*Oversampling*) 139

4.2.4.2.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset* my*Personality* 140

4.2.4.3 Dataset Manual Gathering 141

4.2.4.3.1 Skenario 4 (Percobaan tanpa menggunakan

*Resampling*) 141

4.2.4.3.2 Skenario 5 (Percobaan dengan menggunakan

*Undersampling*) 141

4.2.4.3.3 Skenario 6 (Percobaan dengan menggunakan

*Oversampling*) 142

4.2.4.3.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset Manual Gathering* 143

4.2.4.4 Dataset Gabungan 144

4.2.4.4.1 Skenario 7 (Percobaan tanpa menggunakan

*Resampling*) 144

4.2.4.4.2 Skenario 8 (Percobaan dengan menggunakan

*Undersampling*) 145

4.2.4.4.3 Skenario 9 (Percobaan dengan menggunakan

*Oversampling*) 146

4.2.4.4.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset* Gabungan 147

4.2.4.5 Kesimpulan Implementasi *Deep Learning* 148

4.2.5 Tampilan Layar Aplikasi 151

4.3 Evaluasi 156

4.3.1 Evaluasi Subjektif 156

4.3.2 Evaluasi Objektif 159

4.3.2.1 Sistem Prediksi Kepribadian “*The Big Five Personality*”

Dari Data Twitter 159

4.3.2.2 *Personality Traits Recognition on Social Network -*

*Facebook* 161

4.3.2.3 *Deep Learning-Based Document Modeling for Personality*

*Detection from Text* 164

BAB 5 SIMPULAN DAN SARAN 167

5.1 Simpulan 167

5.2 Saran 168

REFERENSI 169

RIWAYAT HIDUP

LAMPIRAN L1

SURAT NON SURVEI